

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2001, 19:32:01 ; Search time 182.97 Seconds
(without alignments)
7255.390 Million cell updates/sec

Title: US-09-494-297-1

Perfect score: 2274
Sequence: 1 atgaaacaaacaggttc.....gataagaacatgactag 2274

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 678276 seqs, 291890651 residues

Total number of hits satisfying chosen parameters: 1356552

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

N_Geneseq_0401: *
1: /cgnl_8/gcgdata/geneseq/geneseqn/NA1980.DAT: *
2: /cgnl_8/gcgdata/geneseq/geneseqn/NA1981.DAT: *
3: /cgnl_8/gcgdata/geneseq/geneseqn/NA1982.DAT: *
4: /cgnl_8/gcgdata/geneseq/geneseqn/NA1983.DAT: *
5: /cgnl_8/gcgdata/geneseq/geneseqn/NA1984.DAT: *
6: /cgnl_8/gcgdata/geneseq/geneseqn/NA1985.DAT: *
7: /cgnl_8/gcgdata/geneseq/geneseqn/NA1986.DAT: *
8: /cgnl_8/gcgdata/geneseq/geneseqn/NA1987.DAT: *
9: /cgnl_8/gcgdata/geneseq/geneseqn/NA1988.DAT: *
10: /cgnl_8/gcgdata/geneseq/geneseqn/NA1989.DAT: *
11: /cgnl_8/gcgdata/geneseq/geneseqn/NA1990.DAT: *
12: /cgnl_8/gcgdata/geneseq/geneseqn/NA1991.DAT: *
13: /cgnl_8/gcgdata/geneseq/geneseqn/NA1992.DAT: *
14: /cgnl_8/gcgdata/geneseq/geneseqn/NA1993.DAT: *
15: /cgnl_8/gcgdata/geneseq/geneseqn/NA1994.DAT: *
16: /cgnl_8/gcgdata/geneseq/geneseqn/NA1995.DAT: *
17: /cgnl_8/gcgdata/geneseq/geneseqn/NA1996.DAT: *
18: /cgnl_8/gcgdata/geneseq/geneseqn/NA1997.DAT: *
19: /cgnl_8/gcgdata/geneseq/geneseqn/NA1998.DAT: *
20: /cgnl_8/gcgdata/geneseq/geneseqn/NA1999.DAT: *
21: /cgnl_8/gcgdata/geneseq/geneseqn/NA2000.DAT: *
22: /cgnl_8/gcgdata/geneseq/geneseqn/NA2001.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	56.6	2.5	4677	21 A70259	Plasmodium falciparum
2	56	2.5	11922	21 A70187	Plasmodium falciparum
3	53.8	2.4	5940	21 A70105	Plasmodium falciparum
4	51.4	2.3	1998	21 A70212	Plasmodium falciparum
5	51.2	2.3	1527	21 A70121	Plasmodium falciparum
6	51.2	2.3	3399	17 T05868	Chicken leucocytos
7	49.6	2.2	15016	20 X99560	Nucleic acid seque
8	49.6	2.2	3095	11 Q03875	Sequence encoding
9	48.2	2.1	876	14 Q05947	Immunoglobulin bin
10	48.2	2.1	3279	14 Q05946	Sequence encoding
11	48.2	2.1	3279	14 Q51556	Sequence encoding

12	48.2	2.1	20674	21 C58017	Arachidonic acid m
13	47.4	2.1	5897	18 V74631	Staphylococcus aur
14	47.4	2.1	7458	21 A70106	Plasmodium falcipa
15	47.4	2.1	1664976	19 V21209	Methanococcus jann
16	47	2.1	1062	20 X61556	B. burgdorferi ant
17	47	2.1	1132	20 X61555	Nucleic acid seque
18	47	2.1	1785	20 X99656	Nucleic acid seque
19	47	2.1	2418	20 X99561	Nucleic acid seque
20	47	2.1	116277	20 X20249	Borrelia burgdorfe
21	46.4	2.0	5361	18 T78868	P. falciparum live
22	46.4	2.0	6152	18 T78867	P. falciparum live
23	46.2	2.0	2503	15 Q53480	PNP30 xylanase cd
24	46.2	2.0	910715	20 X20248	Borrelia burgdorfe
25	46	2.0	3837	21 A70211	Plasmodium falcipa
26	45.6	2.0	1686	16 Q87587	DNA encoding Leuco
27	45.4	2.0	2841	18 V74488	Staphylococcus aur
28	45.2	2.0	1938	17 T08079	Fibrinogen binding
29	44.8	2.0	997	21 A14996	CDNA encoding a hu
30	44.8	2.0	1220	18 V75229	Staphylococcus aur
31	44.8	2.0	3308	21 A26917	Essential Staphylo
32	44.8	2.0	6042	21 A70199	Plasmodium falcipa
33	44.6	2.0	3642	21 A70180	Plasmodium falcipa
34	44.6	2.0	19124	18 T72882	Plasmodium var-7 g
35	44.6	2.0	19124	21 T98287	Plasmodium var-7 p
36	44.4	2.0	3567	21 A70117	Plasmodium falcipa
37	44.4	2.0	5652	20 X99575	Nucleic acid seque
38	44.4	2.0	14066	20 X99556	Nucleic acid seque
39	44.4	2.0	26811	20 X20253	Borrelia burgdorfe
40	44.2	1.9	6621	21 A70188	Plasmodium falcipa
41	44	1.9	1664976	19 V21209	Methanococcus jann
42	43.8	1.9	5454	21 A70236	Plasmodium falcipa
43	43.6	1.9	4311	21 A70133	Plasmodium falcipa
44	43.4	1.9	732	20 X99612	Nucleic acid seque
45	43.4	1.9	1149	20 X99651	Nucleic acid seque

ALIGNMENTS

RESULT 1	
ID A70259	standard; DNA; 4677 BP.
AC A70259;	
DT 07-NOV-2000	(first entry)
DE Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:392.	
KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine; antimalarial; malaria; protozoacide; infection; insecticide; ds.	
OS Plasmodium falciparum.	
PN W0200025728-A2.	
PD 11-MAY-2000.	
PE 05-NOV-1999;	99WO-US26796.
PR 05-NOV-1998;	98US-0107131.
PA (HOFF/) HOFFMAN S.	
PA (CARU/) CARUCCI D.	
PA (GARD/) GARDNER M.	
PA (VENT/) VENTER J C.	
PI Hoffman S, Carucci D, Gardner M, Venter JC;	
DR WPI; 2000-365347/31.	
PT Proteins encoded by chromosome 2 of the human malarial parasite, Plasmodium falciparum, useful as antimalarial vaccines and in the	

PT diagnosis of *P. falciparum* infection -
 XX
 PS Disclosure: Page 565-566; 577pp; English.
 XX
 CC The present invention describes proteins and their fragments (I) encoded
 CC by chromosome 2 of the human malarial parasite, *Plasmodium falciparum*.
 CC Also described are: (1) nucleotide sequences (II) encoding (I); and (2)
 CC vaccines against *P. falciparum* infection comprising (I) or (II).
 CC (I) and (II) are useful for the development of vaccines against
 CC *P. falciparum* infection. (I) and polyclonal antisera or a monoclonal
 CC antibody raised to immunogens comprising the sequences of (I), are
 CC useful in the detection of infection with *P. falciparum*. Furthermore,
 CC (I) (especially when they are rifins or secreted or membrane proteins)
 CC can aid the identification of drugs to treat or prevent *P. falciparum*
 CC infection, or they can be used to identify drug resistance in
 CC *P. falciparum*. Sequencing of the *Plasmodium* chromosome 2 and the
 CC subsequent identification of proteins encoded by it will help to expand
 CC our understanding of parasite biology, a process hampered by the
 CC complexity of the parasitic lifecycle, and provide new targets for
 CC vaccine and drug development. Parasite resistance to drugs and mosquito
 CC resistance to insecticides have led to a resurgence of malaria in many
 CC parts of the world, and there is a pressing need for vaccines and new
 CC drugs. A70078 to A70287 and B18144 to B18352 represent nucleotide and
 CC protein sequences given in the present invention, but which are not
 CC specifically mentioned within the specification.
 XX
 SQ Sequence 4677 BP; 2106 A; 402 C; 966 G; 1203 T; 0 other;

Query Match 2.5%; Score 56.6; DB 21; Length 4677;
 Best Local Similarity 43.5%; Pred. No. 0.0024;
 Matches 416; Conservative 0; Mismatches 529; Indels 12; Gaps 3;

QY 1084 tttaaggttgaagctggcaagtgatctatctattatgtatgaaacgcatgaaaccc 1143
 DB 3256 tctgatttaaaagatcttgaagaagataataaaagaatlaaaagaacaaactt 3315
 QY 1144 aataaagagatagtaggcttactctagtagaagcatatgatttggaagattagc 1203
 DB 3316 gaaagtgaaattttagaagatttaaaagaatttaaaacttggaaagagatatttgaa 3375
 QY 1204 gtttaactacacaaacttgcanaatttattatgcanaaaataaaatggaagtica 1263
 DB 3376 gagaanaaagaataagaaagatcatttgaaattcgaagaagagtggaanaata 3435
 QY 1264 caggtgtcattgattttagcagatcctaataatctccaccagcttgagatggtggg 1323
 DB 3436 aaagatcttgaagcagataataaaagaagatcattcattagaagttgaaagaaaa 3495
 QY 1324 aaaaacatgactcagacttacaacagagagaagaaataacactatctgcaggtcgt 1383
 DB 3496 aaattagaagaagtagcagatttaaaagaagagtagaa-----cattataaagtggt 3549
 QY 1384 gaccttctaatactatctgtgaacccaagagatacagatcctgaaccttctaanaacat 1443
 DB 3550 gatcgcatataaaaggtttggaagaagatgatttgaagaagatgattttaaagga 3609
 QY 1444 atcaaaaaaataatttggaaggtttacagggaaaagaagaagctatgtgtataggt 1503
 DB 3610 agtatattagacatgttaaaaggagatattgaattgaggtatggaataagaaagttta 3669
 QY 1504 ctacttgacacacaaatgctgcgtctaccacttgaacatatatttccactgaagt 1563
 DB 3670 gaagatgtaaacagcaaaacttggagaagaagttgaatcccttaaaagtglttattcagt 3729
 QY 1564 gctgaattagataaaggtaaaactaaagactatcatgtgttttgagagacatgatatagt 1623
 DB 3730 gc---attagcgtgtagtggaacaaatgaacaaagaaaaaagctcaaacctaa 3786
 QY 1624 acttgagcgtgtctaaacacttctgtgaatagctcaagatgtatcctccaagctta 1683
 DB 3787 ttggaagaagattattataaaagaagtgattaaagaaccaaagaaaaataacacaaa 3846

QY 1684 actgacctgatttcttattccgaataacaataatcatcttatttggaaactcag 1743
 DB 3847 aagaagaagtaaggtttgatttaagaataagaaacaaagaatgaatagtagaagttgaa 3906
 QY 1744 tggcatccagaagatttagttatattatctgtatggaagataaaagaagttatcct 1803
 DB 3907 atgaagaatgaagaatagatagtagaagattagaagaagatgtagaagaatagataagaa 3966
 QY 1804 gtaccatatttaacattgagaanaaagtgactggtttagctgtgacagaactaa 1863
 DB 3967 gataaagttgaagatataagataagatagatagatagatagatagatagattgaa 4026
 QY 1864 gattccatttgaattgaataaaataaataaagaagaatgttcttccaactgtt 1923
 DB 4027 gacaagaatgaagttatagatttaattagtcacaaagaagaaacgattg---aaagatt 4083
 QY 1924 aaaaacgataaacaacactcgaatttaagaatggttaagaacacatttaataacat 1983
 DB 4084 aaagagaaaaaagaanaaattagaaaaaaagttgagaagagtgtagtgccttaaaaa 4143
 QY 1984 ggggaagtttaacacttcaaggtttaccagaaggttatcttacccttgcagaaga 2040
 DB 4144 cactgaagcaagtaatgaaatattgttcaaaaaattgataaagaagttgataaaga 4200

RESULT 2

A70187
 ID A70187 standard; DNA; 11922 BP.

AC A70187;
 DT 07-NOV-2000 (first entry)

XX Plasmodium falciparum chromosome 2 related DNA sequence SRQ ID NO:320.
 DE
 XX
 KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
 KW antimalarial; malaria; protozoacide; infection; insecticide; ds.
 OS
 XX Plasmodium falciparum.
 PN
 XX WO200025728-A2.

PD 11-MAY-2000.
 XX
 PF 05-NOV-1999; 99WO-US26796.
 XX
 PR 05-NOV-1998; 98US-0107131.
 XX

PA (HOFF/) HOFFMAN S.
 PA (CARD/) CARUCCI D.
 PA (GARD/) GARDNER M.
 PA (VENT/) VENTER J C.

XX Hoffman S, Carucci D, Gardner M, Venter JC;

PI WPI; 2000-365347/31.

XX
 DR
 PT Proteins encoded by chromosome 2 of the human malarial parasite,
 PT Plasmodium falciparum, useful as antimalarial vaccines and in the
 PT diagnosis of *P. falciparum* infection -

XX
 PS Disclosure: Page 516-519; 577pp; English.

XX
 CC The present invention describes proteins and their fragments (I) encoded
 CC by chromosome 2 of the human malarial parasite, *Plasmodium falciparum*.
 CC Also described are: (1) nucleotide sequences (II) encoding (I); and (2).
 CC vaccines against *P. falciparum* infection comprising (I) or (II).
 CC (I) and (II) are useful for the development of vaccines against
 CC *P. falciparum* infection. (I) and polyclonal antisera or a monoclonal
 CC antibody raised to immunogens comprising the sequences of (I), are
 CC useful in the detection of infection with *P. falciparum*. Furthermore,
 CC (I) (especially when they are rifins or secreted or membrane proteins)
 CC can aid the identification of drugs to treat or prevent *P. falciparum*

CC infection, or they can be used to identify drug resistance in
 CC P. falciparum. Sequencing of the Plasmodium chromosome 2 and the
 CC subsequent identification of proteins encoded by it will help to expand
 CC our understanding of parasite biology, a process hampered by the
 CC complexity of the parasitic lifecycle, and provide new targets for
 CC vaccine and drug development. Parasite resistance to drugs and mosquito
 CC resistance to insecticides have led to a resurgence of malaria in many
 CC parts of the world, and there is a pressing need for vaccines and new
 CC drugs. A70078 to A70287 and B18144 to B18352 represent nucleotide and
 CC protein sequences given in the present invention, but which are not
 CC specifically mentioned within the specification.

XX Sequence 11922 BP; 5402 A; 948 C; 1343 G; 4229 T; 0 other;

Query Match 2.5%; Score 56; DB 21; Length 11922;
 Best Local Similarity 41.3%; Pred. No. 0.0041;
 Matches 380; Conservative 0; Mismatches 540; Indels 0; Gaps 0;

QY 988 aatgatatggagaaagaaatgaaactatcagatggaacttaacttaactgaatgaat 1047
 DB 10261 aatacaatacaataaataatcttgatcacgataatcattatgatgacataatacaat 10320
 QY 1048 tctccagctggtatgatacgagccacatcactttaagttgaagtcgcaagt 1107
 DB 10321 gatgacataatacatatgacataatacatatgaagcgcataatacatatga 10380
 QY 1108 tatatactatgatgtaaaacagatgtaaaatcccaataagagatagagcccttac 1167
 DB 10381 tataatcataatgatgacataatacatatgatgacataatacatatgatgacataat 10440
 QY 1168 tcagtagaagcataatgattttgaagaatttagcgtttaactacacaaactatgca 1227
 DB 10441 cataatgatgcataataatcctaaatgacagataatcattatgacagataatcag 10500
 QY 1228 aattttatgatcaaaaaataaaatggaagtccaggttgcttctatgctttaatgca 1287
 DB 10501 gataattataatcctcataataataaagaactataaataattatcgatcatgagag 10560
 QY 1288 gatctaaatctccacagactctgaagatggtggaacaaatgacccagacttaca 1347
 DB 10561 gaagatatattatcagaataataattatcacaaagatgattatgttaataaattat 10620
 QY 1348 acagggagaaataaacactcatatgacgctgagccttitaatactatgagaa 1407
 DB 10621 ttcatcaataataataattatcaaaacagctccaagatatagcactattactagt 10680
 QY 1408 ccaagagatccgactctgactcttcttaaaacatacaaaaaaagtaattggaaggt 1467
 DB 10681 tgggggagatataatcaataatccattatctttaataaagaataatataccaaa 10740
 QY 1468 taacagggaaaagaagacagctatgtagttagtgcgttaactgagacacatgctg 1527
 DB 10741 aaaaaaaac 10800
 QY 1528 gctactcagttagcaaatatatttcaactgatagtgctgaattagataagataa 1587
 DB 10801 attatcataatgaagtagatatttcttagaagcattttaaatgftgagggaaat 10860
 QY 1588 aagagactatcgtttttgagacatgagatgacttttagcagttgttaaatcct 1647
 DB 10861 tcagatcgtatattatcttattttaaaaaatcgaaaaaataatgctcagctatctt 10920
 QY 1648 gtgaataacgctaaagatgtaactccacagctactactgactgattcttattccg 1707
 DB 10921 gaagaatccatctatattgtagtactactacacacacacacacacacacacacac 10980
 QY 1708 aatacaataataataatcctcttattggaactcagtgacaccagagaagttagttgat 1767
 DB 10981 aataataataatcctcttataatcttaattgataaaaggtgaacacacattattat 11040
 QY 1768 attatcgttgtagaagataaaaaaagaagtataccctgtaactcattatattgaatgaga 1827

DB 11041 ataatagaatattagaaaaatcattacgaaaaaagaagtattatatcatatagt 11100
 QY 1828 aaaacggtgactggttagtggtagacagacaaagattcatttgaattgaatta 1887
 DB 11101 ccttaagaatattatgatcatcaacaatgtagtacttaataatcagaataaaaaa 11160
 QY 1888 aaaaaataataagaagaatt 1907
 DB 11161 aatacaacaaaaatgaact 11180

RESULT 3

ID A70105 standard; DNA: 5940 BP.
 AC A70105;
 XX
 DT 07-NOV-2000 (first entry)
 XX
 DE Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:238.
 XX
 KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
 XX antimalarial; malaria; protozoacide; infection; insecticide; ds.
 XX
 OS Plasmodium falciparum.
 XX
 PN WO200025728-A2.
 XX
 PD 11-MAY-2000.
 XX
 PF 05-NOV-1999; 99MO-US26796.
 XX
 PR 05-NOV-1998; 980S-0107131.
 XX
 PA (HOFF/) HOFFMAN S.
 PA (CARU/) CARUCCI D.
 PA (GARD/) GARDNER M.
 PA (VENT/) VENTER J C.
 XX
 PI Hoffman S, Carucci D, Gardner M, Venter JC;
 DR WPI; 2000-365347/31.
 XX
 PT Proteins encoded by chromosome 2 of the human malaria parasite,
 PT Plasmodium falciparum, useful as antimalarial vaccines and in the
 PT diagnosis of P.falciparum infection -
 XX
 PS Disclosure: Page 460-462; 577pp; English.
 XX
 CC The present invention describes proteins and their fragments (I) encoded
 CC by chromosome 2 of the human malaria parasite, Plasmodium falciparum.
 CC Also described are: (1) nucleotide sequences (II) encoding (I); and (2)
 CC vaccines against P. falciparum infection comprising (I) or (II).
 CC (I) and (II) are useful for the development of vaccines against
 CC P. falciparum infection. (I) and polyclonal antisera or a monoclonal
 CC antibody raised to immunogens comprising the sequences of (I), are
 CC useful in the detection of infection with P. falciparum. Furthermore,
 CC (I) (especially when they are rifins or secreted or membrane proteins)
 CC can aid the identification of drugs to treat or prevent P. falciparum
 CC infection, or they can be used to identify drug resistance in
 CC P. falciparum. Sequencing of the Plasmodium chromosome 2 and the
 CC subsequent identification of proteins encoded by it will help to expand
 CC our understanding of parasite biology, a process hampered by the
 CC complexity of the parasitic lifecycle, and provide new targets for
 CC vaccine and drug development. Parasite resistance to drugs and mosquito
 CC resistance to insecticides have led to a resurgence of malaria in many
 CC parts of the world, and there is a pressing need for vaccines and new
 CC drugs. A70078 to A70287 and B18144 to B18352 represent nucleotide and
 CC protein sequences given in the present invention, but which are not
 CC specifically mentioned within the specification.
 XX
 SO Sequence 5940 BP; 3106 A; 343 C; 879 G; 1612 T; 0 other;

OY	2178	tcaaaagatcagtctatctacg	2201
Dd	771	aaaagatgaagtacacaataaac	794
RESULT	6		
T05868	ID	T05868 standard; DNA; 3399 BP.	
XX	AC	T05868;	
XX	DT	14-AUG-1996 (first entry)	
XX	DE	Chicken leucocytozoan DNA encoding immunogenic protein for vaccines.	
XX	KW	Chicken leucocytozoan; immunogen; recombinant vaccine; protection;	
XX	RW	Immunisation; vaccination; ss.	
XX	OS	Chicken leucocytozoan.	
XX	FT		
XX	FT	Key Location/Qualifiers	
XX	FT	CDS 1..3399	
XX	FT	/tag= a	
XX	FT	misc.feature 1150..3218	
XX	FT	/tag= b	
XX	FT	/note= "fragment referred to in the claims, for	
XX	FT	use as insert in a recombinant vaccine	
XX	FT	against chicken leucocytozoan disease"	
XX	PN	JP07284392-A.	
XX	PD	31-OCT-1995.	
XX	PF	19-APR-1994; 94JP-0080643.	
XX	PR	19-APR-1994; 94JP-0080643.	
XX	PA	(DOBU-) DOBUTSUO SEIBUTSUGAKUTEKI SEIZAI KYOKAI.	
XX	PA	(KITA) KITASATO KENKYUSHO SH.	
XX	DR	WI: 1996-006311/01.	
XX	DR	P-PSDB; R97866.	
XX	PT	Chicken leucocytozoan immunogenic protein - used in a recombinant	
XX	PT	vaccine against chicken leucocytozoan disease	
XX	PS	Claim 6; Page 6-9; 35pp; Japanese.	
XX	CC	T05868 encodes a chicken leucocytozoan immunogenic protein, this DNA	
XX	CC	or a fragment of it can be used in a recombinant vaccine to immunise	
XX	CC	against chicken leucocytozoan disease. The DNA is used in a vector	
XX	CC	and operatively linked to an expression regulatory sequence as in	
XX	CC	standard practice.	
XX	SQ	Sequence 3399 BP; 1577 A; 508 C; 798 G; 516 T; 0 other;	
Query Match	2.3%	Score 51.2; DB 17; Length 3399;	
Best Local Similarity	41.9%;	Pred. NO. 0.034;	
Matches 445; Conservative	0;	Mismatches 608; Indels 9; Gaps 2;	
OY	1092	tgaagctgcagaagtgatattactattcgatggaaacagatggaanaatcccataaaga	1151
Dd	1542	tgaagaagaataataatttatatagatgatgtaccagaagaatgtcagaagaatcaga	1601
OY	1152	gatgttagagccctaactcagtagaagcatataatgatttgaagaattagcgtttaac	1211
Dd	1602	tgaacaagacatatataacatagaaatagataatgacatacagaagagcatgaaaaagt	1661
OY	1212	taccacaacctatgcaaaattttatattgcaaaaaataaaatggaagtcacagttgt	1271
Dd	1662	aaccaataagaagaagaagaagtagcacatgaaagaataagaaaagaagagcatga	1721

QY	1272	ctatgtcttaataagcagaatcttaaatctccacagagctcttggaagcttggttggaagaacat	1331
Db	1722	agaagatatacacaagaagaagaaagaagaagatatacacatctgaagaataagaagaaga	1731
QY	1332	gactccagactttaaacaacagagaagataaatacactctacatattgcaggtctgtacctt	1391
Db	1782	gcatgaagaagtaatacacaagaagaagaagaagaagatacacatctgaagaagaaga	1841
QY	1332	taataatactgtgaaccacaagagataccgatctgtgacattcttaaacatataaaaa	1451
Db	1842	agaagagcatgaagaagtaatacacaagaagaagaagaagaagaagatacacatgaagaat	1901
QY	1452	agtaattggaagaggtttacagagggaagaagacagactatttgatagtgtgctaaactga	1511
Db	1902	agaaaagaagaagagatgaagaagtaatacacatgaagaagaagaagaagaagatacacatga	1961
QY	1512	gacacaattgcgtgcggtctactcagttacgaatataatttcaactgtatgtgtcyaatt	1571
Db	1962	agaataagaagaagaagatgaagaagatgaagaagtaatacacaatggaagaagaagaat	2021
QY	1572	agataagataaactaaagaactacatgatttttgagacatgaatgtatgtacttgc	1631
Db	2022	acatgaagaagaatagaaaaaagaagaagatgaagaagatacacaatgaagaagaagaaga	2081
QY	1632	agttgtcctaaactctctgtagatcagctccacagaatgaatctccacagctactgacct	1691
Db	2082	agtaacacatagaagaataatagaagaagaagaagatctgaagaagatacatcatgaagaagaaa	2141
QY	1692	tgatttcttcttactccgaataacataatacactccttacttggaaactggtgcac	1751
Db	2142	agaagaagatacacatcatgaagaagaagaagaagaagatctgaagaagt--atacatga	2198
QY	1752	agaagatttggttgatcttcttgatgaagaataaagaagaagtataactgttaacca	1811
Db	2199	agaagaagaagaagaagatacacatgaagaagaagaagaagaagaagatacacatgaagaaga	2256
QY	1812	taatttaacattggaagaacacggtgactggtttagctgtgtgacagaactaaagattcca	1871
Db	2259	agaagaagaagaagtaatacacaagaaga-----aaagaagaagatacatcatgaagaaga	2312
QY	1872	tttgaataattgaattaaataataataagcaagaattgtcttctcaaacctgttaaacaga	1931
Db	2313	aaaagaagaagtaacacacatgaagaagaagaagaagaagatacacatgaagaagaagaaga	2372
QY	1932	taaaacaacactcggaattttaagaatgttaagaaacatataatttaaacatgagggaag	1991
Db	2373	agaagatacacatgaagaagaagaagaagaagaagatacacatgaagaagaagaagaagat	2432
QY	1992	tttaaccttaaaagtttaccagaagaagtttatttcttcctctgtccaagaacacgaattcga	2051
Db	2433	aaccatgaagaagaagaagaagaagtaacacactgagaagaagaagaagaagaagatacacatga	2492
QY	2052	aggtcataaggttaagaatataatgcacagaagaatgacaaatgtctacagtttcaaaaacgg	2111
Db	2493	agaagaagaagaagaagaagatacacatgaagaagaagaagaagaagatacacatgaagaagaaga	2552
QY	2112	aatacaagtgatgaagacactgtcttcttgaataataaaga	2153
Db	2553	aaaagtaacacatgaagaagaagaagaagaagaagtaacatctgaaga	2594
RESULT 7			
X9560/c			
ID	X9560 standard; DNA; 15016 BP.		
XX	X9560;		
XX	05-OCT-1999 (first entry)		
XX	Nucleic acid sequence from U. urealyticum.		
DE	Nucleic acid sequence from U. urealyticum.		
XX	Ureaplasma urealyticum; nucleic acid detection; infection; pathogen;		

KW human urogenital tract; pregnancy; neonatal disease; drug therapy;
 KW suppurative arthritis; ss.
 OS Ureaplasma urealyticum.
 XX MO9939007-A1.
 XX PD 05-AUG-1999.
 XX PF 29-JAN-1999; 99WO-US01972.
 XX PR 30-JAN-1998; 98US-0073189.
 XX PA (UABR-) UAB RES FOUNDED.
 XX PI Cassell GH, Chen EX, Glass JI, Glass JS, Heiner CR;
 XX PI Leikowitz E;
 XX DR WPI: 1999-469343/39.
 XX PT Detection of Ureaplasma urealyticum using novel genes, probes and
 PT primers
 PS Claim 1; Page 48-53; 110pp; English.
 XX The present invention provides methods for the detection and diagnosis
 CC of Ureaplasma urealyticum infection. It provides novel genes (X99501-681)
 CC that can be used as a source of primers and probes for the detection and/
 CC or quantification of U. urealyticum in a biological sample. The probes
 CC that can be used in the method of the invention by forming target:probe
 CC complex is complementary to a region selected from one of the 181
 CC nucleic acid sequences (X99501-681). U. urealyticum is an opportunistic
 CC pathogen of the human urogenital tract that is a significant cause of
 CC adverse pregnancy outcome, neonatal disease, and suppurative arthritis.
 CC As the infections are commonly asymptomatic, it is important to have
 CC specific and sensitive methods for detecting their presence in a patient.
 CC Also, as the pathogen has no current antibiotic directed specifically
 CC against it, it would be advantageous to isolate and detect gene sequences
 CC which are unique to it, and utilise these as a basis for diagnosis of
 CC U. urealyticum infection as well as to develop new and improved drug
 CC therapies. The present invention provides such novel polynucleotide
 CC sequences (X99501-681).
 XX Sequence 15016 BP; 4931 A; 1914 C; 1794 G; 6377 T; 0 other;
 SQ

Query Match 2.2%; Score 50.2; DB 20; Length 15016;
 Best Local Similarity 47.5%; Pred. No. 0.084;
 Matches 182; Conservative 0; Mismatches 198; Indels 3; Gaps 1;

Oy 1781 aagataaaagaagttatctacttaacttaacttaacttgagaagaacggtgactg 1840
 Db 7274 ATGTTAAATATGCTTTATTTATATCAACGCAAAATGCTAAATTTAATTAACATATTACAA 7215
 Oy 1841 gtttagcggtgagcaagcaactaagattccattcgaattgaaattaaataaataaagc 1900
 Db 7214 TTAATATGATCAACAAATTAACATTTCTATTGTGAATTTGATTTGATTAATTTAACTT 7155
 Oy 1901 aagaattgcttcaaacgtttaaacagataaacaacactgaaattaaagatgta 1960
 Db 7154 TAAATCAAGATTATGCTTTTGTGAATTTAGTTGCTAAATAAACCAATTCATGCCGAT 7095
 Oy 1961 aagcaacataatttaaacatcggggaagtttaacacttcaagtttaccagaagttc 2020
 Db 7094 TTGCTTAAATTAATGATATGCCA--ACATCCTTATAGTCTTTTAAAGTACTAATAATCAAG 7038
 Oy 2021 attcttacctgtcaagaagaacatctcgaagcttaaggttaaggttaataccaag 2080
 Db 7037 ATATTAATCAAAAAATTCCTTTAGTTGTGAAGTTAATTAATTAATTAATTAACAG 6978
 Oy 2081 aagtagcaaatgctacagtttcaaaaacagagaataacaagatgatgagcacttgctt 2140
 Db 6977 ATCTTGATTATGATCAAAACCAAGATTAGTTGACCGATTGAATTAAGTTCTATTAT 6918

Oy 2141 aaataaataagagcctgtgtt 2163
 Db 6917 TAAACAATAGTACATTCGTGTT 6895

RESULT 8
 ID Q03875 standard; DNA; 3095 BP.
 AC Q03875;
 XX 24-AUG-1990 (first entry)
 DE Sequence encoding carboxylic terminal part of native GLURP.
 XX Plasmodium falciparum; antigen; malaria; vaccine; GLURP:ss
 KW Plasmodium falciparum.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 1..2352
 FT /tag= a
 FT /product=GLURP
 XX
 XX MO9022811-A.
 XX PD 22-MAR-1990.
 XX PF 18-SEP-1989; 89WO-0000218.
 XX PR 03-MAR-1989; 89US-0218885.
 XX PR 03-MAR-1989; 89DK-0005191.
 XX PA (STAT-) STATENS SERUMINST.
 XX PI Dziegiel M, Borre M, Jepsen S, Vuust J, Rieneck K, Wind A;
 XX PI Jakobsen PH;
 XX DR WPI: 1990-115998/15.
 XX P-PSDB; R05804.
 XX PT Polypeptide(s) derived from Plasmodium falciparum antigen - used in
 PT vaccines and in production of antibodies, for diagnosis and
 PT therapy of malaria.
 XX
 PS Disclosure: Fig 7; 108pp; English.
 XX An open reading frame of 2349 bps extends from the 5' terminal end of the
 CC insert to a "TAA" stop codon. It is longest ORF found in the sequence.
 CC Sequence displays some of the characteristics of other malaria nucleic
 CC acid sequences: tandemly repeated motifs, high AT content and a
 CC corresponding preference for codons containing these bases, and a high
 CC content of codons for glutamate. Three major repetitive sequences are:
 CC one motif from bp 34 to bp 156 is repeated from bp 289 to bp 411; another
 CC motif from bp 477 to bp 521 is repeated tandemly twice from bp 522 to bp
 CC 566 and from bp 567 to bp 611; a third motif from bp 1174 to bp 1233 is
 CC repeated tandemly 11 times. This last repetitive region consists of 360bp
 CC repeats differing only in 3 bases GAT coding for aspartate. This region
 CC is flanked to the 5' terminal od a degenerated 60 bp repeat. GC content
 CC of the coding part of the insert is on average 30%, and of the non-coding
 CC 3' terminal 11%.

XX Sequence 3095 BP; 1443 A; 300 C; 491 G; 861 T; 0 other;
 SQ

Query Match 2.2%; Score 49.6; DB 11; Length 3095;
 Best Local Similarity 42.9%; Pred. No. 0.075;
 Matches 320; Conservative 0; Mismatches 414; Indels 12; Gaps 1;

Oy 1453 gtaattgagaaggttcaagaagaagaagcaagctatgagttagtgctactagag 1512
 Db 1057 gaatttgagagagtttccctcgaaacaaatcaaaataacgaattccaagaattaatgaa 1116

```

QY 1513 acacaatgctgctgactcagttacgaatattatattcactgtagtgcgaatta 1572
DB 1117 gatgataaaggtgacacatcagcagtaaatagtagaagaagaatactccagaa 1176
QY 1573 gataagataaactaaagactacatcagtttttgagacatgaaatgtagtacttagca 1632
DB 1177 gatgataaataatgtaaaagttagacatgtaaatagtagaagttgagaataattctagcagaa 1236
QY 1633 gttgctaaatcctctgtagataacgctcagaatagtagtaattcctccacagctaacctt 1692
DB 1237 gataaaaatgaaaaaggtcacaatgtaaatagtagaaggttgagaagaattctaccagaaagat 1296
QY 1693 gattctttctccgaataacaataatcatcctctatttggaactcagtgccatca 1752
DB 1297 gataaaaatgaaaaaggtcacaatgtaaatagtagaaggttgagaagaattctaccagaaagat 1356
QY 1753 gaagatttagttgatatattcgtatggaagataaaaaagaagtatacctgtaacctat 1812
DB 1357 aaaaatgaaaaaggtcacaatgtaaatagtagaaggttgagaagaattctaccagaaagataaa 1416
QY 1813 aattacatctgaaaaaaacggctgactggttttagctggtgacagaactaaagatttccat 1872
DB 1417 aatgaaaaaggtgacacatgtaaatgtagaaggttgagaagaattctaccagaaagataaaat 1476
QY 1873 ttggaattgtaataaaaaataaataagcaagaattgctttctcacaactgtttaaacagat 1932
DB 1477 gaaaaaaggtcacaatgtaaatagtagaaggttgagaagaattctaccagaaagataaaatgaa 1536
QY 1933 aaaaacaaacctcgaaatttaagaatgtaaaagcaaccataatttaaacatggygaagat 1992
DB 1537 aaagtccaacatgtaaatagtagaaggttgagaagaattctaccagaaagataaaatgaaaaa 1596
QY 1993 ttaacacttaaggtttaccagaaggttatcttactctgccaagaagaacagattctgaa 2052
DB 1597 ggtcaacatgtaaatagtagaaggttgagaagaattctaccagaagaagataaaatgaaaaa 1656
QY 2053 ggcata-----taaggttaaggttaataagccaagaagtagcaaatgctacagtt 2100
DB 1657 ggtcaacatgtaaatagtagaaggttgagaagaattctaccagaagaagataaaatgaaaaagtt 1716
QY 2101 tcaaaaacaggaataacaagaatgtagacacatgcttttgaaataataaagaagcctggtt 2160
DB 1717 caacatgtaaatagtagaaggttgagaagaattctaccagaagaagataaaatgaaaaagttcaa 1776
QY 2161 gttcctcagaggtgtagtcaaaagat 2186
DB 1777 catgaaatagtagaaggttgagaagaat 1802

RESULT 9
Q50947
ID Q50947 standard: cDNA; 876 BP.
XX
XX
AC Q50947;
XX
XX
DT 18-MAY-1994 (first entry)
XX
DE Immunoglobulin binding protein gene derived from protein L gene.
XX
KW Peptide; immunoglobulin; binding; analysis; purification; ELISA;
KW enzyme linked immunosorbent assay; ss.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
XX repeat_unit 1..96
XX FT /*tag- a
XX FT /note- "Repeat units are not adjacent, repetitions
XX FT of this sequence are not 100% homologous and
XX FT begin at nucleotide positions 214, 427 and
XX FT repeat_unit 97..210 649"

```

```

FT FT /*tag- b
FT FT /note- "Repeat units are not adjacent, repetitions
FT FT of this sequence are not 100% homologous and
FT FT begin at nucleotide positions 313, 535 and
FT FT 757"
MO9322439-A.
XX
XX PN
XX
XX PD 11-NOV-1993.
XX
XX PE 07-MAY-1993; 93WO-GB00950.
XX
XX PR 07-MAY-1992; 92GB-0009804.
XX
XX PR 24-DEC-1992; 92GB-0026928.
XX
XX PA (PUBL-) PUBLIC HEALTH LAB SERVICE BOARD.
XX
XX PI Atkinson A, Dugleby CJ, Murphy JP, Trowern AR;
XX
XX PI WPI: 1993-368798/46.
XX
XX DR P-PSDB; R42204.
XX
XX PT New immunoglobulin binding proteins derived from protein L -
XX PT which bind immunoglobulin kappa light chains but not albumin or
XX PT cell walls
XX
XX PS Claim 13; Figure 2; 28pp; English.
XX
XX CC The synthetic immunoglobulin binding proteins derived from protein
XX CC L comprise repeated sequences from protein L which bind
XX CC immunoglobulin kappa light chains. They can be used in protein
XX CC analysis, purification procedures and other biochemical processes e.
XX CC g. ELISA. The synthetic molecules are of particular advantage if
XX CC they are free of regions in protein L which exhibit albumin and cell
XX CC wall binding.
XX
XX SQ Sequence 876 BP; 402 A; 141 C; 162 G; 171 T; 0 other;

Query Match 2.1%; Score 48.2; DB 14; Length 876;
Best Local Similarity 46.7%; Pred. No. 0.11;
Matches 189; Conservative 0; Mismatches 213; Indels 3; Gaps 1;

QY 1061 atagatcgaagacacatcactttaaggttgaagctggaaggtgtatataattatg 1120
DB 362 aaatgctggaatataacagcagacttagaagatggtggaacacatcaatcaattatg 421
QY 1121 atggaacacagattgaaatcccaataaagaatagtagagccttactcagtagaagcat 1180
DB 422 ctggaagaagaacacacagaaacacagaaagacccaagaagaaggttacaatacaagta 481
QY 1181 ataagtttttgaagaatttagcgttttaactacacaaactatgcaaaattttattatg 1240
DB 482 acttaacttttgacagatggaagaatatacaacagcagaatccaaggaacatttgaagaag 541
QY 1241 caaaaataaataatggaagttcacaggttgcatttctattgcttaatgacagacttaaatctc 1300
DB 542 caaca---gcaaaagcttaatgcttaatgcaaatcaatattagcaaaaagaatggygaatata 598
QY 1301 caccagactctggaagatggtggaacaaatgtagctcagaacttacaacagagaagtaa 1360
DB 599 cagaagacttagaagatggtggaacaaatcaatcaatcaatatttgcgtggaagaacac 658
QY 1361 aatacactcatattgacaggtcgtgacctttaaataatctggaacccaagataccg 1420
DB 659 cagaacacccaagaagacccaagaagaagttacaatcaaatgtaacttattcttgcag 718
QY 1421 atctgaccttcttaaacatatacaaaaagtaatttgaagag 1465
DB 719 atggaacacccaagaacagcagaatccaagaagaacatttgaagaag 763

RESULT 10

```

Q50946
 ID Q50946 standard; cDNA; 3279 BP.
 AC Q50946;
 XX 18-MAY-1994 (first entry)
 DT
 XX Sequence encoding protein L.
 DE
 XX Peptide; immunoglobulin; binding; analysis; purification; ELISA;
 KW enzyme linked immunosorbent assay; ss.
 OS
 XX Peptococcus magnus.
 XX
 FH Location/Qualifiers
 FT 103..3185
 FT /*tag= a
 FT /product= Protein L.
 FT repeat_unit
 FT 490..573
 FT /*tag= b
 FT /note= "Repeat units are not adjacent, repetitions
 FT of this sequence are not 100% homologous and
 FT begin at nucleotide positions 673 and 856"
 FT repeat_unit
 FT 574..672
 FT /*tag= c
 FT /note= "Repeat units are not adjacent, repetitions
 FT of this sequence are not 100% homologous and
 FT begin at nucleotide position 757"
 FT repeat_unit
 FT 949..1044
 FT /*tag= d
 FT /note= "Repeat units are not adjacent, repetitions
 FT of this sequence are not 100% homologous and
 FT begin at nucleotide positions 1162, 1375
 FT and 1597"
 FT repeat_unit
 FT 1045..1158
 FT /*tag= e
 FT /note= "Repeat units are not adjacent, repetitions
 FT of this sequence are not 100% homologous and
 FT begin at nucleotide positions 1261, 1483
 FT and 1705"
 FT repeat_unit
 FT 1822..1938
 FT /*tag= f
 FT /note= "Repeat units are not adjacent, repetitions
 FT of this sequence are not 100% homologous and
 FT begin at nucleotide positions 2347, 2545 and
 FT 2731"
 FT repeat_unit
 FT 2914..2934
 FT /*tag= g
 FT /note= "Repeat units are adjacent, repetitions
 FT of this sequence are not 100% homologous and
 FT begin at nucleotide positions 2935, 2953,
 FT 2968, 2986, 3001, 3019 and 3034"
 PN W09322439-A.
 XX 11-NOV-1993.
 PD
 XX 07-MAY-1993; 93MO-GB00950.
 PE
 XX 07-MAY-1992; 92GB-0009804.
 PR 24-DEC-1992; 92GB-0026928.
 XX
 XX (PUBL-) PUBLIC HEALTH LAB SERVICE BOARD.
 PA
 PI Atkinson A, Dugleby CJ, Murphy JP, Trowern AR;
 XX
 DR WPI: 1993-368798/46.
 DR P-PSDB; R42203.
 XX
 XX New immunoglobulin binding proteins derived from protein L -
 PT which bind immunoglobulin kappa light chains but not albumin or
 PT cell walls

PS Disclosure; Figure 1; 28pp; English.
 XX
 CC The synthetic immunoglobulin binding proteins derived from protein
 CC L comprise repeated sequences from protein L which bind
 CC immunoglobulin kappa light chains. They can be used in protein
 CC analysis, purification procedures and other biochemical processes e.
 CC g. ELISA. The synthetic molecules are of particular advantage if
 CC they are free of regions in protein L which exhibit albumin and cell
 CC wall binding (The repeat regions commencing at nucleotide numbers
 CC 1045, 1261, 1483 and 1705).
 CC
 XX Sequence 3279 BP; 1505 A; 481 C; 625 G; 668 T; 0 other;
 SO
 Query Match 2.1%; Score 48.2; DB 14; Length 3279;
 Best Local Similarity 46.7%; Pred. NO. 0.16;
 Matches 189; Conservative 0; Mismatches 213; Indels 3; Gaps 1;
 QY 1061 atagatcgagagcacttctttaaagtgtagcgtgcaagtgtagtatttg 1120
 DB 1310 aaatgycgaatatcacagactgagaagatggtgaaacacacacataaattg 1369
 QY 1121 atggaacacagattgaaatcccaataaagagatagtagccttactcagtagaagcat 1180
 DB 1370 ctggaataaagaacacccagaaacacagaaacacaaagaagaagtacataaagtta 1429
 QY 1181 ataagtatttgaagaatttgcgttttaactacacaaactatgcataatttattatg 1240
 DB 1430 acttaattcttcagatgagaagatcacaaacagcagaattccaagaacatttgaagag 1489
 QY 1241 caaaaataaataatggaagtcacagtgctctattgctttaaagcagatctaaatctc 1300
 DB 1490 caaca--gcaaaagcttatgtcttatgcaaaactattagcaaaagaatgycgaatata 1546
 QY 1301 caccagactcgaagatggtggaataacatgactccagacttaccacagagaagtaa 1360
 DB 1547 cagcagacttgaagaatggtggaacacacataaatttgcgtggaagaagaacac 1606
 QY 1361 aatacactcatattgtaggtgtagcctttaaataatactgtgaacccaagagatacgc 1420
 DB 1607 cagaacacccagaagaacacaaagaagaagtacatcaagaattcaacttaactcttcag 1666
 QY 1421 atcctgaacttctttaaatacatcaaaaagtaattgagaag 1465
 DB 1667 atggaataacacaaacgcagaaattcaagaagacatttgaagaag 1711
 RESULT 11
 ID Q51556 standard; cDNA; 3279 BP.
 AC Q51556;
 XX
 DT 18-MAY-1994 (first entry)
 DE
 XX Sequence encoding protein L.
 DE
 KW Protein; immunoglobulin; binding; immobilisation; light chains;
 KW antibodies; diagnosis; pharmaceutical; ss.
 XX
 OS Peptococcus magnus.
 XX
 FH Location/Qualifiers
 FT 103..3185
 FT /*tag= a
 FT /product= Protein L.
 FT repeat_unit
 FT 490..573
 FT /*tag= b
 FT /note= "Repeat units are not adjacent, repetitions
 FT of this sequence are not 100% homologous and
 FT begin at nucleotide positions 673 and 856"
 FT repeat_unit
 FT 574..672
 FT /*tag= c

FT	/note=	"Repeat units are not adjacent, repetitions of this sequence are not 100% homologous and begin at nucleotide position 75"
FT		
FT	repeat_unit	949..1044
FT	/*tag=	d
FT	/note=	"Repeat units are not adjacent, repetitions of this sequence are not 100% homologous and begin at nucleotide positions 1162, 1375 and 1597"
FT		
FT	repeat_unit	1045..1158
FT	/*tag=	e
FT	/note=	"Repeat units are not adjacent, repetitions of this sequence are not 100% homologous and begin at nucleotide positions 1261, 1483 and 1705"
FT		
FT	repeat_unit	1822..1938
FT	/*tag=	f
FT	/note=	"Repeat units are not adjacent, repetitions of this sequence are not 100% homologous and begin at nucleotide positions 2347 and 2545"
FT		
FT	repeat_unit	1939..2007
FT	/*tag=	g
FT	/note=	"Repeat units are not adjacent, repetitions of this sequence are not 100% homologous and begin at nucleotide positions 2479, 2665 and 2851"
FT		
FT	repeat_unit	2035..2094
FT	/*tag=	h
FT	/note=	"Repeat units are not adjacent, repetitions of this sequence are not 100% homologous and begin at nucleotide position 2209"
FT		
FT	repeat_unit	2095..2208
FT	/*tag=	i
FT	/note=	"Repeat units are not adjacent, repetitions of this sequence are not 100% homologous and begin at nucleotide positions 2269"
FT		
FT	repeat_unit	2914..2934
FT	/*tag=	j
FT	/note=	"Repeat units are adjacent, repetitions of this sequence are not 100% homologous and begin at nucleotide positions 2935, 2953, 2968, 2986, 3001, 3019 and 3034"
XX		
PN	W09322438-A.	
XX		
PD	11-NOV-1993.	
XX		
PF	07-MAY-1993;	93WO-GB00949.
XX		
PR	07-MAY-1992;	92GB-0009804.
XX		
PA	(PUBL-) PUBLIC HEALTH LAB SERVICE BOARD.	
XX		
PI	Atkinson A, Dugleby CJ, Murphy JP, Trowern AR;	
DR	WP1, 1993-368797/46.	
XX	P-PSDB; R43699.	
XX		
PT	Immunoglobulin binding polypeptide, protein L - used for produ.	
PT	of pharmaceuticals and for immobilising antibodies e.g. on	
PT	columns, in diagnostic tests and in assays	
XX		
XS	Disclosure: Figure 1; 29pp; English.	
XX		
CC	Protein L forms a complex with immunoglobulin kappa light chain.	
CC	Purified protein can be used as a reagent for immobilising	
CC	antibodies e.g. on columns, in diagnostic tests and in assays. It	
CC	may also be used in the production of pharmaceuticals.	
XQ	Sequence 3279 BP; 1505 A; 480 C; 626 G; 668 T; 0 other;	

[illegible]

RESULT	12	
ID	C58017	standard; DNA; 20674 BP.
AC	C58017;	
DT	25-JAN-2001	(first entry)
DE	Arachidonic acid metabolism related genomic biallelic marker #651.	
XX		
XX	Human; biallelic marker; arachidonic acid metabolism; genotyping;	
KW	detection; hybridisation; phenotype; haplotype; SNP; polymorphic base;	
KW	single nucleotide polymorphism; hybridisation assay; sequencing assay;	
KW	specific amplification assay; identification; ERBW; 12-LO-RBM;	
XX	eicosanoid-related biallelic marker; 12-LO-related biallelic marker; ds	
OS	Homo sapiens.	
XX		
PN	WO200047771-A2.	
XX		
PD	17-AUG-2000.	
XX		
PF	11-FEB-2000; 2000WO-1B00184.	
XX		
PR	12-FEB-1999; 99US-0119917.	
PR	23-MAR-1999; 99US-0275267.	
PR	07-MAY-1999; 99US-0133200.	
XX		
PA	(GEST) GENSET.	
XX		
PI	Blumenfeld M, Bouquelieret L, Chumakov I;	
XX		
DR	WPI; 2000-571891/53.	
XX		
PT	Novel biallelic markers useful for detecting conditions and genotypes	
XX	associated with arachidonic acid metabolism -	
XX		
PS	Claim 67; Page 790-796; 802pp; English.	
XX		
CC	The present invention describes polynucleotides including biallelic markers derived from genes involved in arachidonic acid metabolism and	

CC from genomic regions flanking those genes. Methods from the present
CC invention may be used to select individuals for clinical trials and
CC predict responses to treatment with drugs. The polynucleotides may be
CC used in hybridisation assays, sequencing assays and specific
CC amplification assays for identifying an ascotensoid-related allelic
CC marker (RBM) or 12-10-related allelic marker, and for amplifying a
CC segment of nucleotides containing an RBM. The polynucleotides are
CC useful in diagnostic kits. The markers may be used to detect conditions
CC and genotypes associated with arachidonic acid metabolism. C57367 to
CC C58018 and B24019 and B24020 represent sequences used in the
CC exemplification of the present invention.
CC N.B. Polymorphic bases (single nucleotide polymorphisms also known as
CC SNPs) in the polynucleotide sequences from the present invention have
CC been given as their corresponding degenerate bases e.g. a polymorphic
CC base of C or T has been given as Y.

SQ Sequence 20674 BP; 5400 A; 5170 C; 4995 G; 5062 T; 47 other;

Query Match	2.1%	Score 48.2	DB 21	Length 20674
Best Local Similarity	45.9%	Pred. No. 0.25		
Matches 240; Conservative	0	Mismatches 278	Indels 5	Gaps 2

QY	1608	agacatgcatgatgacatttagcaggtgcgaacacccgtgagatgcgtccaaagatag	1667
Db	11084	atattttaatatttttaattcaataaattttaatttaatatataaataaataatttaatt	11143
QY	1668	taatccctcacagctaactgcagcttgatttccttaatccgaatacaataaatcaatc	1727
Db	11144	taatatataatataaattttaattttaataataataatataaataatataaatttaatt	11203
QY	1728	tctattggaactcgaatgcgcacccaagaattgtgtatattatctgtatgaagatcaa	1787
Db	11204	taaaatttaaatcttaattttaatacaaacatttaaatctttaatttaatatataattat	11263
QY	1788	aaaagaagtataactgcgttaactc--taattacaattggaaaaacgctgactggttt	1844
Db	11264	ttaatttaatatataatttaacttaatttaatttaatttaatttaatttaa--cttaatttaatt	11321
QY	1845	agcgcgggacagcaactaaagtctccacttttgaattgaatttaaaaaataataagcaaga	1904
Db	11322	aatttaatatataatataaattttaatatattttaatttaatttaatatataatttaattta	11381
QY	1905	atgctttctcaaacctgttaaacacagataaacaacacccctgaatttaagaatgcygaagc	1964
Db	11382	atttaaatatttaatttaatttaattttaattttaattgcttcgaataatttaatttaatttaa	11441
QY	1965	aaccataatttaaacacatggyggaaagtcttaacacctcaaggtttcacagaaggtatcc	2024
Db	11442	tatttttaatttaatatataatataaattttaattttaataatataatataatttaatat	11501
QY	2025	ttacctgtcacaagaacacagatctcgaagctgtatgaagtttaagtttaatgcagaagagt	2084
Db	11502	taatttaaatatttaatatataaattttaatagcttaagtaagaacatcttttaatatataaaaaag	11561
QY	2085	agcaaatgctacagttcttcaaaacacaggaatacaagttatgag	2127
Db	11562	accatgaccttttttcaatagcttaaggaagaacgcaggaagaag	11604

RESULT	13
V74631	
ID	V74631 standard; DNA; 5897 BP.
XX	
AC	V74631;
XX	
DT	16-MAR-1999 (first entry)
XX	
DE	Staphylococcus aureus contig SEQ ID #320.
XX	
KM	Computer readable medium; vaccine; S.aureus infection; immunodetection;
KM	cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW	skin infection; surgical wound infection; scalded skin syndrome;

KW toxic shock syndrome; ds.

OS Staphylococcus aureus.

FH	key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

misc_feature

```
ET / *lag = a
ET / notes = "++"
```

13

are included to maintain the nucleotide numbering

13

```

ET      misc_feature      2221. . 2280
ET      /#330- b

```

```

E1 /- lay = 0
ET /note= "tr

```

the sequence listing in the specification. They

LE

91
1007

```

FT      misc_feature      4021..4080
FT

```

FT /note= "these bases represent a line of missing text in

ET

ET
E

air

ET	mic feature	ET	mic feature
5827	5880	91	91

FT
E A mmsc_recurate
/ *tag= d

ET

Et

LE 13

XX 13

PN EP786519-A2.

XX

PD 30-JUL-1997.

XX
PF 07-JAN-1997.

XX 5
C
135
135

PR 05-JAN-1996;

XX
XX

Sequence 5897 BP; 2046 A; 852 C; 1114 G; 1642 T; 243 other;

Search completed: June 6, 2001, 21:48:17
Job time: 8176 sec

```

PD 26-FEB-1998.
XX
XX 22-AUG-1997; 97WO-US14900.
XX
PR 22-AUG-1996; 96US-0024428.
XX
PA (GENO-) INST GENOMIC RES.
PA (UNIT ) UNIV ILLINOIS FOUND.
PA (UJDO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
PI Bult CJ, Smith HO, Venter JC, White OR, Woese CR;
XX WPI; 1998-169145/15.
DR
XX Complete genome sequence of methano-genic archaeon, Methanococcus
PT jannaschii - useful in identification of M. jannaschii genome
PT fragment
XX
PS Claim 13: Page 152-585; 614pp; English.
XX
XX The present sequence represents the complete 1.66-megabase pair genome
CC sequence of the Methanococcus jannaschii circular chromosome. The
CC present invention describes M. jannaschii open reading frames from the
CC genome sequence. The invention also describes a computer based system
CC for identifying fragments of the M. jannaschii genome that are
CC homologous to target nucleotide sequences, comprising: (a) data storage
CC means comprising the nucleotide sequence of the 1664976, 58407 or 16550
CC bp sequence (see V21209, V21210 and V21211), or a nucleotide sequence at
CC least 99.9% identical to it; (b) search means for comparing a target
CC sequence to the nucleotide sequence of the data storage means to
CC identify a homologous sequence, and (c) retrieval means for obtaining
CC the homologous sequence. The method, which is based on whole genome
CC random sequencing of an autotrophic archaeon M. jannaschii, the genome
CC of which consists of 3 physically distinct elements, a large circular
CC chromosome (the 1664976 bp sequence given in V21209), a large circular
CC extra-chromosomal element (the 58407 bp sequence given in V21210), and a
CC small circular extra-chromosomal element (the 16550 bp sequence given in
CC V21211), can be used in the identification of M. jannaschii genome
CC fragment.
XX
SQ Sequence 1664976 BP; 568133 A; 264649 C; 258701 G; 573392 T; 101 other;

Query Match 2.1%; Score 47.4; DB 19; Length 1664976;
Best Local Similarity 46.3%; Pred. No. 1.2;
Matches 156; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

QY 1778 tggaaataaaagaagtataccgttaactataatattaaatgagaagaacggtga 1837
DB 201070 ttgataataataagaattatgtgcttaactgaaaaattagaagagcttaaaatatta 201129
QY 1838 ctggttagctgtgacagaaactaaagattccatttgaattgaattaaataataa 1897
DB 201130 aagatggcttgaaagatttaataatataatgcaacttaagatttagcaatagataaca 201189
QY 1898 agcaagaattgcttctcaaatgtttaaacaagataaacaacacctogaatttaagatg 1957
DB 201190 ttaagaggaagtataataagaagatattgaattaccctaacaacaatttagag 201249
QY 1958 gtaagaacaccatttaataaacaatggggaaggttaaacattcaagtttaccagaag 2017
DB 201250 ttaataaggaatttaattgatatagaagaagaatattcctactaaccaaaaacttgatg 201309
QY 2018 gtattcttactctgtcaagaagaacagattctgaagctataaggttaagtttaagtcacc 2077
DB 201310 aaataaactacaatgaagaagaacataaaaaataaagaagccttatgaaaataagaagac 201369
QY 2078 aagaagtaagcaaatgctacagtttcaaaaacaggaat 2114
DB 201370 aagaactgtataacgtagaagaacaaaaacagaat 201406

```

